Genet Resour Crop Evol (2013) 60:1997–2015
DOI 10.1007/s10722-013-9968-1

RESEARCH ARTICLE

Genetic diversity and origin of cultivated potatoes based on plastid microsatellite polymorphism

Tatjana Gavrilenko · Olga Antonova · Anna Shuvalova · Ekaterina Krylova · Natalia Alpatyeva · David M. Spooner · Lubov Novikova

Received: 24 August 2012 / Accepted: 16 January 2013 / Published online: 8 May 2013

Abstract The origin of cultivated potatoes has remaining questions. In this study, 237 accessions of all cultivated species and 155 accessions of wild species closely related to cultivated potatoes, including their putative ancestors, were analyzed using 15 plastid microsatellites (SSRs) to investigate genetic diversity and their relationships with the wild species. We here used polymorphic plastid SSRs we developed from potato plastid genome sequences as well as already known plastid SSR markers. All 15 loci were polymorphic and identified a total of 127 haplotypes. Dramatic decreases in levels of genetic diversity were revealed in landraces in comparison with wild ancestor species. The plastid SSR results showed a decrease in haplotype number and diversity from Peru to both north and south. Phylogenetic analysis revealed two distinct groups. One of them, group A, contained the majority of accessions of cultivated species of the *Solanum tuberosum* Andigenum group including all accessions of cultivated diploid and triploid cytotypes of this group (*S. chaucha*, *S. phureja*, and *S. stenotomum* by a former taxonomic system) and most of tetraploid accessions of the *S. tuberosum* Andigenum group (*S. tuberosum* subsp. *andigenum*), and the majority of accessions of wild ancestors from the northern members of the *S. brevicaule* complex. Another group B comprised most of the wild species accessions and almost exclusively hybrid cultivated species which have introgressed plastid genomes from the other wild gene pools. Lack of clustering of traditional cultivated species (as used above) support a revised group classification of *S. tuberosum*.

Electronic supplementary material The online version of this article (doi:10.1007/s10722-013-9968-1) contains supplementary material, which is available to authorized users.

T. Gavrilenko (✉) · O. Antonova · A. Shuvalova · E. Krylova · N. Alpatyeva · L. Novikova
N. I. Vavilov Institute of Plant Industry, Bolshaya Morskaya Street, 42-44, 190000 Saint Petersburg, Russia
e-mail: tatjana9972@yandex.ru

O. Antonova
e-mail: olgaant326@qip.ru

A. Shuvalova
e-mail: islamshina@pochta.ru

E. Krylova
e-mail: katrin@newmail.ru

N. Alpatyeva
e-mail: n.alpatyeva@vir.nw.ru

L. Novikova
e-mail: lubov.novikova@mail.ru

T. Gavrilenko
Saint Petersburg State University, Saint Petersburg, Russia

D. M. Spooner
USDA, Agricultural Research Service, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI, USA
e-mail: david.spooner@ars.usda.gov

Keywords Cultivated potatoes · Genetic variation · Plastid microsatellites (SSRs) · Phylogeny · *Solanum* · Taxonomy

Introduction

Solanum L. section *Petota* includes about 100 wild and cultivated species (Spooner 2009) (see Table 1 for authors of taxonomic names). All are tuber-bearing, and the world-wide cultivated potato *S. tuberosum* L. represents the third economically most important world food crop. Potato species are distributed from the southwestern United States to south-central Chile and adjacent Argentina (Hijmans et al. 2002) and exist as a polyploid series ranging from diploid to hexaploid wild species. Cultivated potato landraces are restricted to South America, from western Venezuela to south-central Chile. The domestication process involved the selection of wild potato genotypes producing larger tubers with lower glycoalkaloid content, shorter stolons, and attractive colors of tuber skin and tuber flesh. These genotypes were the subject of human and natural selection. Subsequent to selection, some of them further hybridized and/or underwent ploidy changes that today have resulted in a diverse genepool of thousands diverse landraces. The center of diversity of the group of cultivated species was placed in the central Andes where diploid, triploid, tetraploid, and pentaploid landraces grow; in Chile there are predominately tetraploid landraces.

Differentiation within cultivated species complex

Cultivated potatoes have been classified by different characters and wide variation exists within most of the species. Only Chilean landraces have a distinct habitat, adapted to the lowlands of south-central Chile, in contrast to all other landraces that occur naturally in the mid to high elevations of the Andes Mountains (Spooner et al. 2010), with the intervening range between Andean and Chilean potatoes of 560 km separated by the Atacama desert. Physiological characters also have differentiated these landraces, with the Chilean potatoes able to produce tubers under the long days of central coastal Chile and the other landraces adapted to the shorter day length. Some diploid accessions of *S. tuberosum* Andigenum group, formerly classified as *S. phureja*, were thought to be most common to lower and warmer regions of the Andes and to possess tubers lacking dormancy. Three highland frost resistant potato species (*S. ajanhuiri*, *S. curtilobum*, *S. juzepczukii*) are cultivated at altitudes between 3,600 and 4,300 m with temperatures during the growing period between 6° and 14 °C, sometimes dropping to −5 °C (Brown 1999; Huanco 1991). Only highland frost resistant potatoes could be differentiated from the remaining Andean landraces based on phenetic analysis of numerous morphological characters (Huamán and Spooner 2002; Gavrilenko et al. 2010). Biochemical characters divide landraces into bitter potatoes (*S. curtilobum*, *S. juzepczukii*, and some cultivars of *S. ajanhuiri*) whose tubers can be used for food after freeze-drying to remove high level of glycoalkaloids, and other cultivated species whose tubers can be eaten without special processing (Brown 1999). Diploid members of the *S. tuberosum* Andigenum group are self-incompatible outbreeders which are readily crossable with each other and with many of the wild species relatives (see review in Spooner et al. 2010); tetraploid members of this group are self- and cross-compatible whereas Chilean tetraploids are often characterized by pollen sterility controlled by nuclear genes and putative cytoplasmic factors (Grun 1973; Grun et al. 1977); cultivated triploids and pentaploids are mostly sterile (Hawkes 1990).

DNA methods provide additional data about genetic differentiation of cultivated potatoes. Thus, Hosaka and collaborators, using plastid DNA analyses, divided cultivated potatoes into groups having several main plastid types designated as A, C, S, T, and W types (Hosaka et al. 1984; Hosaka 1986, 1995; Hosaka and Hanneman 1988a, b). None of these types were species-specific, however cultivated species were characterized by their different frequencies. Thus, T type plastid DNA (corresponding to the presence of a 241 bp deletion in the *ndhC/trnV* intergenic spacer region; Kawagoe and Kihuta 1991) was found only in tetraploid cultivated species, predominantly in Chilean, and rarely in Andean tetraploid landraces (Hosaka 1995, 2004). Also, nuclear microsatellites or simple sequence repeats (SSRs) support the genetic difference between most tetraploid Andean and Chilean landraces (Raker and Spooner 2002; Spooner et al. 2007b; Gavrilenko et al. 2010); their separation was also shown by nuclear RFLPs (Sukhotu et al. 2004). Nuclear SSRs separate *S. ajanhuiri*, *S. curtilobum*, and *S. juzepczukii* from the

Table 1 Species analyzed in plastid SSR analysis

Name of species according to Spooner et al. (2007a, b) and Ovchinnikova et al. (2011)	Number of accessions analyzed	Abbreviation	Name of accessions according to Hawkes (1990)	Ploidy level	Number of accessions analyzed
Cultivated potato species					
S. ajanhuiri Juz. et Bukasov	3	ajh	*S. ajanhuiri* Juz. et Bukasov	2x	3
S. juzepczukii Bukasov	12	juz	*S. juzepczukii* Bukasov	3x	12
S. curtilobum Juz. et Bukasov	10	cur	*S. curtilobum* Juz. et Bukasov	5x	10
S. tuberosum L.					
Andigenum Group diploids	104	Adg2x	*S. phureja* Juz. et Bukasov	2x	40
- /-		Adg2x	*S. stenotomum* Juz. et Bukasov	2x	64
S. tuberosum L.					
Andigenum Group triploids	27	Adg3x	*S. chaucha* Juz. et Bukasov	3x	27
S. tuberosum L.					
Andigenum Group tetraploids	47	Adg4x	*S. tuberosum* subsp. *andigenum* Hawkes	4x	47
S. tuberosum L.					
Chilotanum Group	34	Chl	*S. tuberosum* subsp. *tuberosum* Juz. et Bukasov	4x	34
				Total	237

Name of species according to the Solanaceae Source website	Number of accessions analyzed	Abbreviation	Name of series and species according to Hawkes (1990)	Number of accessions analyzed
Wild species			Series Acaulia Bitter	
S. acaule Bitter	16	acl	*S. acaule* Bitter	16
S. albicans (Ochoa) Ochoa	6	alb	*S. albicans* (Ochoa) Ochoa	6
			Series Maglia Bitter	
S. maglia Schltdl.	3	mag	*S. maglia* Schltdl.	3
			Series Megistacroloba Cárd. et Hawkes	
S. boliviense Dunal	11	blv	*S. megistacrolobum* Bitter	5
-/-		blv	*S. toralapanum* Cárd. et Hawkes	6
			Series Yungasensa Correll	
S. berthaultii Hawkes	15	ber	*S. tarijense* Hawkes	12
			Series Tuberosa (wild) (Rydb.) Hawkes	
-/-		ber	*S. berthaultii* Hawkes	3
			Southern members of the *S. brevicaule* complex:	

Table 1 continued

Name of species according to the Solanaceae Source website	Number of accessions analyzed	Abbreviation	Name of series and species according to Hawkes (1990)	Number of accessions analyzed
S. brevicaule Bitter	74	brc	*S. brevicaule* Bitter	6
-/-		brc	*S. gourlayi* Hawkes	10
-/-		brc	*S. hondelmannii* Hawkes et Hjert.	11
-/-		brc	*S. leptophyes* Bitter	10
-/-		brc	*S. oplocense* Hawkes	8
-/-		brc	*S. sparsipilum* (Bitter) Juz. et Bukasov	13
-/-		brc	*S. spegazzinii* Bitter	16
			Northern members of the *S. brevicaule* complex:	
S. candolleanum Berthault.	19	cnd	*S. bukasovii* Juz.	7
-/-		cnd	*S. canasense* Hawkes	2
-/-		cnd	*S. multidissectum* Hawkes	10
S. vernei Bitter et Wittm.	8	vrn	*S. vernei* Bitter et Wittm.	8
			Outgroup:	
			Series Pinnatisecta (Rydb.) Hawkes	
S. pinnatisectum Dun.		pnt	*S. pinnatisectum* Dunal	1
S. tarnii Hawkes et Hjert.		trn	*S. tarnii* Hawkes et Hjert	1
			Far outgroup (not a member of section *Petota*)	
S. nigrum L.		ngr	*S. nigrum* L.	1
			Total	155

remaining cultivated species (Spooner et al. 2007b; Gavrilenko et al. 2010).

Different characters mentioned above were used for taxonomic treatments of cultivated potato species which has been the subject of much debate (see review in Ovchinnikova et al. 2011). We here use the latest classification system for group of cultivated species of Spooner et al. (2007b) and Ovchinnikova et al. (2011). For wild species taxonomy we follow the Solanaceae Source website (http://www.nhm.ac.uk/research-curation/research/projects/solanaceaesource/). At the same time we also link here the classification system of Hawkes (1990) to be related to literature (Table 1, Supplemental file 1).

Origin of cultivated potatoes

The wild ancestors and center(s) of origin of cultivated potatoes are still in dispute. Two main hypotheses on the origin of cultivated potatoes were proposed.

(1) A multiple-origin hypothesis was first proposed by Russian scholars (Juzepczuk and Bukasov 1929; Bukasov 1933, 1939, 1966, 1978; Juzepczuk 1937), postulating that the current distribution of cultivated potatoes reflects their geographic origin from different wild gene pools. Each of the Andean cultivated diploid species was domesticated from different diploid wild species distributed in Peru, Bolivia, Colombia; subsequent polyploidization of diploid cultivated species led to formation tri- and tetraploid landraces in Andes (except bitter potatoes).

Following Darwin (1845), Russian taxonomists believed that Chilean landraces evolved in situ from putative Chilean wild species (Juzepczuk and Bukasov 1929). Later, the domestic origin of Chilean landraces from wild species *S. maglia* (endemic to Chile and a single valley in Argentina adjacent to Chile) was proposed by Ugent et al. (1987). This hypothesis found partial support from nuclear SSR data, grouping *S. maglia* with Chilean landraces (Spooner et al. 2012).

(2) A restricted origin hypothesis was developed by Salaman (1946) and elaborated by Simmonds (1964) and Hawkes (1956, 1990) who proposed that domestication took place in a relatively restricted geographical region in southern Peru and northern Bolivia from single ancestor as the '*Solanum brevicaule* complex' (Ugent 1970) or one diploid wild species *S. leptophyes* (Hawkes 1990), followed by polyploidization and interspecific hybridization with other wild species. AFLP analysis of landraces and wild species supports a monophyletic origin of all cultivated potatoes (except bitter potatoes) in the area of southern Peru from northern wild species in the *S. brevicaule* complex (Spooner et al. 2005). These results also were supported by previous data from RAPDs and low-copy nuclear RFLPs (Miller and Spooner 1999) and morphology (Van den Berg et al. 1998).

Authors of the restricted origin hypothesis proposed that the Chilean landraces were transported from those originally formed in the Andes, not originated in situ as proposed by Russian workers, with adaptation to long days occurring in coastal Chile. Grun (1990) proposed that Chilean cultivated potatoes originated from interspecific hybridization between an unidentified wild species as the maternal parent and tetraploid Andean landraces as the paternal parent. Hosaka's group identified such a wild species as *S. tarijense* (=*S. berthaultii*; Spooner et al. 2007a), distributed from central Bolivia to northern Argentina. According to their concept these interspecific hybrids were transferred from the Andes to the southern regions to Chile, based on finding T type plastid DNA in ~5 % of tetraploid Andean landraces, in ~10 % of wild species *S. berthaultii*, and in ~90 % of Chilean landraces (Hosaka 2002, 2003, 2004).

Both hypotheses are concordant regarding the origin of bitter potatoes growing in the Altiplano in that they proposed their hybrid origin after potato was domesticated in the highlands of Peru and Bolivia. Based on morphological, ecogeographical, physiological, and ploidy level characters Bukasov (1939) proposed that triploid *S. juzepczukii* originated through natural hybridization between cultivated diploids and frost resistant tetraploid wild species *S. acaule*; subsequent natural crosses between *S. juzepczukii* (forming unreduced gametes) and Andean tetraploid landraces led to formation pentaploid *S. curtilobum*. Experimental conformation of Bukasov's prediction was provided through re-synthesis of these cultivated species by Hawkes (1962) and Schmiediche et al. (1982). The same approach was applied for *S. ajanhuiri* to support its hybrid origin from the wild progenitor *S. boliviense* (Huamán et al. 1982). Later, hybrid origins of frost-resistant bitter potatoes was conformed using DNA markers of AFLPs (Kardolus 1998), nuclear RFLPs (Sukhotu et al. 2004), nuclear SSRs (Spooner et al. 2007b; Gavrilenko et al. 2010) and DNA sequence data (Rodríguez et al. 2010). Plastid genome studies are

important in identifying parentages in hybrid speciation and domestication. Unfortunately, only few accessions of bitter potatoes were analyzed using plastid DNA markers, so there are gaps in maternal origin data.

The purpose of our study was to investigate, through a comprehensive set of all cultivated and related wild species, the use of plastid SSRs, and the 241 bp plastid deletion marker, cultivated potato diversity, origin, and taxonomy. Most of screened 392 accessions were studied previously in phenetic analysis of morphological characters and in nuclear SSR analysis (Gavrilenko et al. 2010), that here allowed us to compare our present plastid data to our prior morphological and nuclear SSR results.

Materials and methods

Accessions examined

We examined 392 accessions of cultivated and wild potato species, mostly from the experimental subset of the VIR potato collection, recently characterized by nuclear SSRs and morphology (Gavrilenko et al. 2010), but also including accessions from four other genebanks listed in Supplemental file 1. These included 237 accessions of all cultivated species taxa and 155 accessions of eight wild species, and three more distantly related outgroups in section *Petota* (*S. pinnatisectum*, *S. tarnii*), and a non-tuber-bearing far outgroup, *S. nigrum* (Table 1, Supplemental file 1). Each accession was represented by a single genotype.

DNA extraction

Total genomic DNA was extracted according to a modified CTAB method (Saghai-Maroof et al. 1984) from plants grown in a field plot at the VIR experimental station, Pushkin, Saint Petersburg, Russia. Approximately 100 mg of leaf tissue was ground in liquid nitrogen, resuspended in 600 µl of pre-warmed CTAB extraction buffer (100 mM Tris HCl, pH 8.0; 1.2 M NaCl; 20 mM EDTA pH = 8.0; 2 % CTAB; 1 % polyvinylpyrrolidone and 0.2 % β-mercaptoethanol), and incubated at 65 °C for 2 h. After cooling to room temperature, 100 µl of RNAseA solution (10 mg/ml) was added, and the mixture was left to stand at room temperature for 20 min. Then, 100 µl 12 % sodium metabisulphite was added (so that the final concentration of $Na_2S_2O_5$ was 1.5 %) and then 1,000 µl chloroform/isoamyl alcohol solution (24:1). The mixture was gently mixed by inverting the tubes for 40 min. The phases were separated in a microfuge at 3,500 g for 10 min at room temperature. A total of 1,400 µl cold ethanol was added to the approximately 600 µl of aqueous phase and left to precipitate overnight at 20 °C. DNA was collected by centrifugation (5,000 g; 4 °C; 15 min), washed three times with 80 % ethanol, and dissolved in 500 µl of TE buffer (10 mM Tris–HCl, pH = 8,0; 1 mM EDTA, pH = 8.0).

SSR development and PCR analysis

We developed 12 SSR primer pairs (designated as STCP) using the program "Primer 3″ (http://primer3.sourceforge.net/webif.php) to amplify mononucleotide repeats in the single copy regions of the plastid genome (Supplemental file 2), based on the potato plastid genome sequence available in GenBank (accession No. DQ231562.1 (http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=nucleotide&dopt=GenBank&RID=4R02AKGV016&log%24=nucltop&blast_rank=26&list_uids=82754608) (Chung et al. 2006). STCP primer pairs targeted new SSRs of the plastid genome which were not tested before using potato species.

In addition, we included three SSR primer pairs which have been previously designed using a tobacco plastid DNA sequence (NTCP6, NTCP12, NTCP14) by Bryan et al. (1999) and Provan et al. (1999). Among the 15 SSRs, 14 were located in intergenic spacer regions and one (STCP11) in an *ndhA* intron (Supplemental file 3). PCR reactions were performed in a total volume of 12 µl containing 10 ng DNA template, 1 × PCR reaction buffer (Metabion, http://www.metabion.com) with 2.5 mM $MgCl_2$, 0.2 mM of each dNTP (Metabion), 0.2 µM of each primer (forward and reverse), 25 pM of 700 or 800 IRDye-labeled M13 forward primer and 0.5 U Taq polymerase (Metabion). Reactions were carried out using the following program with touchdown profile: 3 min at 94 °C, 5 cycles [45 s at 94 °C, 1 min at T°m + 5° with decrease 1 °C per cycle, 1 min at 72 °C], 5 cycles [45 s at 94 °C, 1 min at T°m, 1 min at 72 °C], 30 cycles [45 s at 94 °C, 45 s at T°m, 1 min at 72 °C] with a final extension step of 5 min at 72 °C.

PCR products were separated by electrophoresis on a LiCor 4300S DNA analyzer system. Size calibration was performed with the molecular weight ladder 50–350 bp using IRD700/IRD 800 (LiCor) by using

SAGA Generation 2 software. Every sample was analyzed at least twice to ensure genotype reproducibility. Alleles were scored as present (1) or absent (0).

Outgroup accessions of *S. pinnatisectum* and *S. tarnii* gave the amplification products with all 15 plastid SSRs used here (although very different in size compared with subset of cultivated and related to them wild species), whereas 13 plastid SSR primers did not amplify DNA from the accession of more distantly related species *S. nigrum*, even after many attempts. We calculated these alleles as real zero alleles because these primer pairs gave amplification products with the other accessions.

Detection of the T type of plastid DNA

We also used a pair of primers H1 (Hosaka 2002) which detects the presence or absence of the 241 bp deletion in the *ndh*C/*trn*V intergenic region corresponding to T type of plastid DNA. PCR reactions and separation of amplified products by electrophoresis were performed according to Hosaka (2002).

DNA cloning and sequencing

We generated DNA sequences from PCR products of primer pairs STCP2 and STCP11 which were of a much smaller size (possible deletions) in some accessions in comparison with the other fragments of the same loci from the majority of accessions. For a control we used DNA of *S. tuberosum* VIR 3446, whose DNA also was amplified with primer pairs STCP2 and STCP11. Amplified PCR products were cut from the gel and ligated into the vector pAL-TA (Evrogen, Moscow) according to the protocol of Evrogen (http://evrogen.ru/kit-user-manuals/pAL-TA.pdf). The ligation products were transformed into competent cells of *E. coli* DH5L. Sequencing was performed on 24-capillary 3500xL Genetic Analyzer (Applied Biosystems, http://www.appliedbiosystems.com) in the Center for Collective Use 'Genomic Technology and Cell Biology' (ARRIAM RAAS, arriam.spb.ru/rus/ckp.html). Alignment of nucleotide sequences and their analysis were conducted using the program MEGA version 4 (Tamura et al. 2007).

Data analyses

The polymorphic information content (PIC) of plastid SSRs was calculated as $PIC = 1-\sum(p_i^2)$, where p_i is the frequency of the p_i allele detected in all accessions (Nei 1973). Haplotypes were defined as unique combinations of alleles at all 15 plastid SSR loci examined. Numbers of haplotypes were estimated for each cultivated species. The haplotype diversity was calculated using the 'haplotype diversity index' (h) as: $h = 1-\sum(g_i^2)$, where g_i is the frequency of the i-th haplotype in each taxa or in other case—the number of haplotypes detected in the same country.

Genetic diversity analysis was performed with the program DARwin5 (http://darwin.cirad.fr/darwin). A dissimilarity matrix was calculated by using the Dice coefficient. The dendrogram was built by using the Neighbor-Joining method (NJ) with bootstrap analyses using 1,000 replicates.

Correspondence between the present plastid data and prior nuclear SSR and morphology data obtained in Gavrilenko et al. (2010) was tested using correlation coefficients of 248 accessions shared between both data sets, and the significance of this correlation coefficient was determined using a Mantel test with 1,000 permutations in NTSYS-pc.

Results

Allelic composition of the plastid SSR haplotypes

Both STCP and NTCP primer pairs generally produced amplification products of the expected size range based on plastid genome sequences of potato (Chung et al. 2006). NTCP primer pairs were previously tested on a wide range of potato landraces by Sukhotu et al. (2004, 2005) and Sukhotu and Hosaka (2006). The size range of fragments generated by NTCP6 and NTCP14 primer pairs with our subset were similar to Sukhotu et al. (2005). However, for the locus *rps2/rpoC2* (NTCP12 primer pair) accessions from our study had the allele sizes within the range of 124–129 bp (Supplemental files 3, 4), whereas Sukhotu et al. (2005) reported allele sizes within the 235–239 bp that is very different from the expected size of a NTCP12 amplified fragment based on DQ231562.1 sequence of *S. tuberosum* 'Desiree' (Chung et al. 2006). In total, 93 alleles were detected in 15 loci. The number of alleles per locus ranged from 3 to 12 with an average of 6.2 and the PIC value ranged from 0.097 to 0.894 (Supplemental file 3).

The 15 plastid SSR loci produced 127 plastid haplotypes for the 392 accessions. Twenty-eight of these haplotypes (#I–#XXVIII; Table 2; Supplemental file 1)

consist of more than one accession per haplotype and included 293 of the 392 accessions. Supplemental file 4 documents the allelic composition of these 28 haplotypes. The remaining 99 haplotypes were unique, each represented by only one accession (Table 2, Supplemental file 1).

We here document three new deletions in the intergenic spacer *rps16/trnQ* (primer pair STCP2) and in the intronic region of *ndhA* (primer pair STCP11). Primer pair STCP2 amplified the fragments in the range of 122–126 bp in most accessions, and one accession of *S. brevicaule* PI 558062 had the smaller fragment of 114 bp and three accessions of *S. candolleanum* (having haplotype #XVIII) had the smallest fragment of 89 bp (Supplemental file 4). Similarly, most accessions of our subset had alleles of locus STCP11 in the range of 135–137 bp, whereas accession PI 472988 of *S. brevicaule* possesses allele STCP11_113. To confirm the presence of deletions in these accessions the corresponding PCR products were cloned and sequenced. Sequencing of the SSR-containing PCR products of these regions confirmed the expectations. Thus, in *S. brevicaule* PI 558062 we detected a 10 bp deletion in region *rps16/trnQ* and the mononucleotide repeat of this accession is two bp shorter than the control accession (Supplemental file 5a). A 36 bp deletion was discovered at the same *rps16/trnQ* region outside the SSR repeat in all three Peruvian accessions of *S. candolleanum* (Supplemental file 5a) specific for haplotype #XVIII. In the intronic region of *ndhA* (STCP11 primer pair) another 21 bp deletion occurs in *S. brevicaule* (PI 472988), and one G > T substitution (transversion) occurs in this region (Supplemental file 5b).

Likewise, haplotypes #I and #VIII are easily identified by visual examination of bands in gels using the NTCP6 primer pairs (region *rps16/trnQ*) that distinguish the shorter size amplicons. Most accessions have alleles of this locus *rps16/trnQ* in the range of 171–176 bp and all accessions with haplotypes #I and #VIII have much smaller fragments of 127 bp and 126 bp respectively (Supplemental file 4). Allele NTCP6_127 was described earlier by Hosaka (2003) as having a 48 bp deletion in *rps16/trnQ* region.

Plastid SSR haplotype distribution in cultivated and wild species

A total of 237 cultivated species accessions grouped in 21 haplotypes, with four most common haplotypes #I-

Table 2 Taxonomic distribution of the plastid SSR haplotypes

Haplotype #	Number of accessions with this haplotype	Species(number of accessions)[a]
I	122	Adg2x(86), Adg3x(14), Adg4x(7), ajh(1), ber(1), brc(1), cnd(1), cur(10), juz(1)
II	48	Adg2x(14), Adg3x(11), Adg4x(19), Chl(1), mag(3)
III	36	Adg4x(3), ber(3), Chl(30)
IV	14	acl(2), ajh(1), juz(11)
V	5	Adg4x(3), Chl(2)
VI	5	Adg3x(1), Adg4x(4)
VII	5	acl(1), alb(4)
VIII	4	Adg4x(4)
IX	4	brc(4)
X	4	acl(4)
XI	4	brc(4)
XII	4	brc(3), cnd(1)
XIII	3	brc(3)
XIV	3	brc(3)
XV	3	brc(3)
XVI	3	brc(1), vrn(2)
XVII	3	ber(3)
XVIII	3	cnd(3)
XIX	2	acl(2)
XX	2	cnd(2)
XXI	2	cnd(2)
XXII	2	cnd(2)
XXIII	2	brc(2)
XXIV	2	brc(1), vrn(1)
XXV	2	vrn(2)
XXVI	2	ber(2)
XXVII	2	brc(2)
XXVIII	2	brc(2)
Unique haplotypes	14	Unique haplotypes in cultivated species:
		Adg2x(4), Adg3x(1), Adg4x(7), ajh(1), Chl(1)
	82	Unique haplotypes in wild species:
		acl(7), alb(2), ber(6), blv(11), brc(45), cnd(8), vrn(3)
		Outgroup:
	3	ngr(1), pnt(1), trn(1)
Total 127	392	

All accessions of haplotype III have the 241-bp plastid deletion (T-type DNA) and all accessions with haplotype XVI have a 65-bp insertion in *ndh*C/*trn*V intergenic region; all other accessions of our subset lack these deletion or insertions

[a] See Table 1 for cultivated and wild species codes

#IV containing 88 % (209 of 237) of the cultivated species accessions. There was much greater diversity in the wild species than the cultivated species; 155 wild species accessions contained 110 haplotypes, 85 of which were unique (Table 2).

The most frequent haplotype #I grouped 122 (31 %) of 392 accessions of our subset; 119 of which were cultivated species (Table 2). This haplotype was detected in half (119 of 237) of the cultivated species accessions and in most (147 of 178 or 83 %) accessions of the *S. tuberosum* Andigenum group including *S. tuberosum* Andigenum Group diploids (86 of 104 accessions), *S. tuberosum* Andigenum Group triploids (14 of 27), and *S. tuberosum* Andigenum Group tetraploids (7 of 47). Haplotype #I was detected in all ten accessions of pentaploid species *S. curtilobum*, in one of three accessions of *S. ajanhuiri*, in one of 12 accessions of *S. juzepczukii* and in three wild species accessions of *S. brevicaule*, *S. berthaultii*, *S. candolleanum*, (Table 2, Supplemental file 1).

Haplotype #II contained 48 (12 %) of the 392 accessions of our subset, containing 45 (19 %) cultivated species accessions, combining 44 accessions of *S. tuberosum* Andigenum Group, and one of the 34 accessions of *S. tuberosum* Chilotanum Group, and all three accessions of the wild species *S. maglia*.

Haplotype #III contained 36 (9 %) accessions of our subset, mostly including *S. tuberosum* Chilotanum Group (30 of the 34 accessions), along with three accessions of *S. tuberosum* Andigenum Group 4x and with three accessions of wild species *S. berthaultii* (Table 2, Supplemental file 1).

Haplotype #IV contained most (11 of 12) accessions of *S. juzepczukii*, one of *S. ajanhuiri*, and two accessions of wild species *S. acaule* (Table 2).

Haplotypes #V, #VI and #VIII each contained small number of accessions of exclusively cultivated species, mainly tetraploid Andean landraces. The remaining 14 cultivated species accessions were represented by unique haplotypes (Table 2, Supplemental file 1).

The remaining 144 wild species accessions were combined in 106 different haplotypes—59 of them had small in number haplotypes (#VII, #IX–#XXVIII) and 85 accessions possess unique haplotypes (Table 2).

In summary, only two taxa were uniform, containing only a single plastid DNA haplotype (one cultivated species *S. curtilobum*, and one wild species *S. maglia*) with haplotype diversity index (*h*) zero, while other taxa contained two (*S. juzepczukii*) to 57 (*S. brevicaule*) haplotypes (Table 2). Among the cultivated species, *S. tuberosum* Andigenum Group 4x was the most heterogeneous (Table 2, Supplemental file 6) possessing 13 different haplotypes (with the highest *h* value of 0.789). *S. ajanhuiri* (diploid) also has a high *h* value ($h = 0.667$), but it must be interpreted with caution because this taxon had only three accessions. The low *h* values occurred in *S. juzepczukii* ($h = 0.153$) and in *S. tuberosum* Chilotanum group ($h = 0.216$) as well as in *S. tuberosum* Andigenum Group 2x ($h = 0.296$) (Supplemental file 6).

Geographical distribution of plastid SSRs in cultivated species

The distribution of plastid haplotypes in cultivated species accessions analyzed in our study spanned from +10 to −45 latitude (Supplemental files 7, 8). Haplotypes #I and #II are widely overlapping in their distributions throughout much of Andean South America. For example, haplotype #I is found from western Venezuela and northern Colombia south to northern Argentina in regions between +10 and −25 latitude. It is prevalent in the northern (+5–0 latitude) and in central Andes (−10 – −15 latitude), rare in northern Argentina, and absent in Chile (Supplemental file 7). In the north, between 0 and −5 latitude, haplotype #I occurs at lowest frequency, and no other haplotypes were found in this region (Supplemental file 7). Haplotype #II occurs almost in all distribution areas, distributed from northern Colombia south to northern Argentina and one accession of *S. tuberosum* Chilotanum Group with haplotype #II detected in Chile. Haplotype #II is frequent in the central Andes at latitude −10 – −15, and distributed to the north and south of this region at low frequencies (Supplemental file 7). In contrast, haplotype #III is mainly distributed in Chile where minimal haplotype diversity was found. Our data show that *h* values decrease from the central Andes to both the north and the south (Chile) (Supplemental file 8). All seven plastid haplotypes (#I–#VI and #VIII shared among cultivated species with different frequencies) occur in southern Peru between −15 and −10 latitude, and most of the unique haplotypes of cultivated species accessions also were found in this region (Supplemental file 8). At the same time, the diversity index of haplotypes (*h* value) in Peruvian accessions was not the highest because 83 % of them shared two main

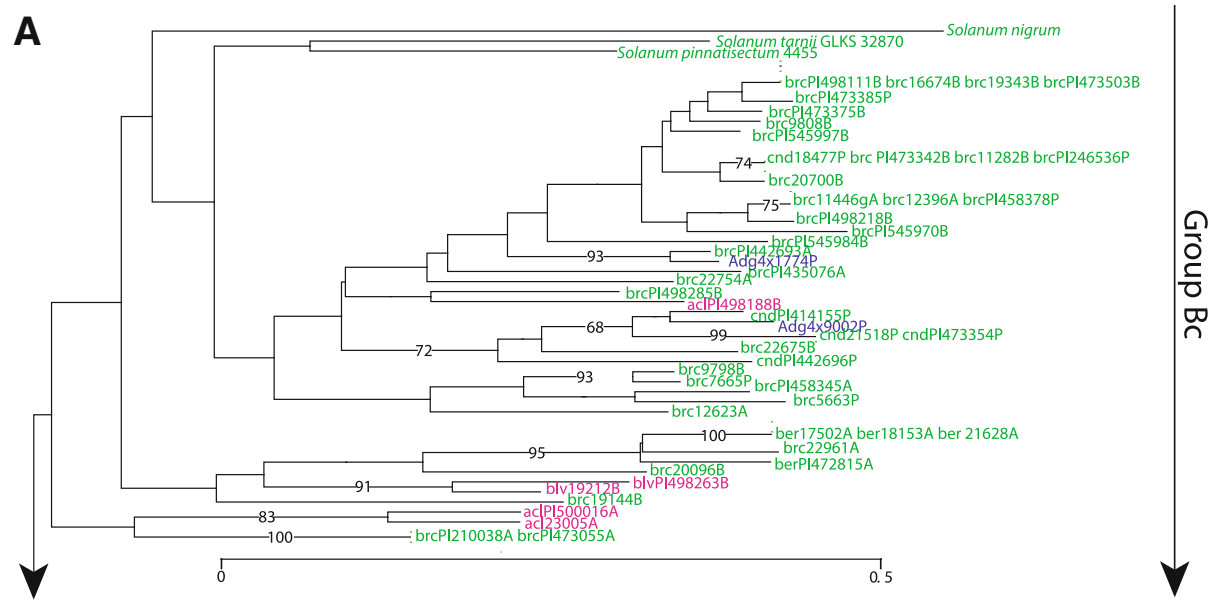

Fig. 1 Accessions of cultivated and wild potatoes examined in this study; species codes are as below, followed by the accession numbers (Supplemental file 1) and the first or first two letters of the countries of origin from Argentina (A), Bolivia(B), Chile (Ch), Colombia (C), Ecuador (E), Peru (P), and Venezuela (V).*Pink*—*S. ajanhuiri* (ajh), *S. juzepczukii* (juz), *S. curtilobum* (cur) and related wild species *S. acaule* (acl), *S. albicans* (alb), *S. boliviense* (blv). *Blue*—*S. tuberosum* Andigenum Group 2x, 3x, and 4x genotypes. *Black*—*S. tuberosum* Chilotanum Group. *Green*—Wild species (outgroups), and closely related to cultivated potato species: *S. berthaultii* (ber), *S. brevicaule* (brc), *S. candolleanum* (cnd), *S. maglia* (mag), *S. vernei* (vrn)

haplotypes #I and #II (Supplemental file 8). The highest value of haplotype diversity occurs in northern Argentina ($h = 0.691$), although we had the lowest number of Argentinean landraces (n = 9), precluding firm conclusions about centers of plastid DNA diversity. Our data suggest, however, that plastid haplotype diversity in cultivated potatoes is high in the central Andes, lower in the northern Andes (Ecuador, Columbia) and significantly lower in lowland Chile (Supplemental file 8).

Neighbor joining results

Accessions of cultivated and wild species are divided into two main groups (labeled Group A, Group B, based on the criteria of separating most cultivated from wild species accessions; Fig. 1). We also labeled the seven (#I–#VI, #VIII) cultivated species haplotypes within groups A and B that are significant for our results.

Group A

Group A includes most of cultivated species accessions (181 of 237 or 76 %) possessing haplotypes #I, #II, #VI, #VIII and six unique haplotypes (Fig. 1). Two haplotypes (#I, #VIII) are very similar and their representatives have the same allelic composition in all analyzed SSR loci except of locus NTCP6 in which the difference in allele size between haplotypes #I and #VIII was only 1 bp (Supplemental file 4). This possibly indicates a recent origin of haplotype #VIII that evolved from ancestors having haplotype #I. Group A contains not only the majority of the cultivated species accessions, but also a minority of wild species accessions, 22 (14 %) of 155, including 14 from Peru, one from Bolivia, four from Argentina and three from Chile (Fig. 1). There are most representatives of *S. candolleanum* (12 of 19), a species now combining many separate species formerly referred to as the northern members of the *S. brevicaule* complex (Table 1). However, seven accessions of *S. candolleanum* are also found in group B, and group A also contains four of 74 accessions of *S. brevicaule*, a species now combining many separate species formerly referred to as the southern members of the *S. brevicaule* complex from Bolivia and Argentina (Table 1). Group A also includes the wild species *S. berthaultii* (one of 15 accessions),

Fig. 1 continued

S. boliviense (two of 11), and all 3 accessions of *S. maglia* (Fig. 1).

Group B

Group B includes most of wild species accessions, 133 (86 %) of 155, and relatively low number of cultivated species accessions, 56 (24 %) of 237, which are representatives of the plastid haplotypes #III, #IV, #V, and eight unique cultivated species haplotypes. Group B divides in three groups Ba, Bb and Bc (Fig. 1).

Group Ba

Group Ba contains most representatives of wild species: *S. acaule* (13 of 16), *S. albicans* (5 of 6), *S. boliviense* (5 of 11 accessions) (Fig. 1; Table 2). Two extra accessions of *S. boliviense* fall as a separate minor group adjacent to the group Ba (Fig. 1). Most accessions of two cultivated species *S. juzepczukii* (11 of 12, all having haplotype #IV) and *S. ajanhuiri* (two of three) also occur in group Ba.

Group Bb

The cultivated species occurring in group Bb contain half (42 of 81) of the tetraploid landraces of *S. tuberosum* of our subset, including 33 *S. tuberosum* Chilotanum Group and nine accessions of Andigenum Group 4x. Thirty-three landraces (30 Chilean and 3 Andean) have plastid haplotype #III (100 % bootstrap) and five landraces have haplotype #V (99 % bootstrap) (Fig. 1). Four landraces (one Chilean and 3 Andean) each with a unique haplotype that grouped adjacent the haplotype #V. Group Bb also contains one of 104 accessions of *S. tuberosum* Andigenum Group 2x with a unique haplotype (Fig. 1).

Of 58 wild species accessions in the group Bb, only four were from Peru, 23 from Bolivia, and 31 from Argentina. Half (37 of 74) of the accessions of *S. brevicaule* and two of 19 *S. candolleanum* accessions (former southern and northern members of *S. brevicaule* complex) are in group Bb. In addition, ten of 15 accessions of *S. berthaultii* occur in the group Bb, including three accessions of this species having haplotype #III (Fig. 1).

Group Bc

This Group contains only two cultivated species accessions of *S. tuberosum* Andigenum Group 4x with unique haplotypes. It contains mostly wild species accessions from different countries: Peru (9), Bolivia (22) and Argentina (16) and from different species: *S. acaule* (3 of 16), *S. boliviense* (2 of 11), *S. brevicaule* (33 of 74), *S. berthaultii* (4 of 11), *S. candolleanum* (5 of 19) (Fig. 1). Neighbor joining clearly separates the outgroups (the non-tuber-bearing *S. nigrum*, and the distantly related Mexican diploid potato species *S. pinnatisectum* and *S. tarnii*) from the remaining wild potato species in the group Bc.

Correspondence between plastid and nuclear SSR and morphological data

Correspondence between the 248 accessions analyzed in common in the present plastid SSR study with the nuclear SSR study of Gavrilenko et al. (2010) was low (0.48, using the Mantel matrix-correspondence test in NTSYS-pc). We also examined correspondence of plastid data obtained here and of the morphological phenetic analysis of Gavrilenko et al. (2010). The correspondence was low (0.08, using the Mantel matrix-correspondence test in NTSYS-pc) for 265 common accessions used in both studies. However, the high similarity of the current plastid results and morphology data has been obtained for *S. tuberosum* Andigenum Group 3x (as *S. chaucha*), based on dividing its accessions on two main morphotypes with plastid haplotypes #I and #II (Supplemental file 9).

Polymorphism in plastid region *ndhC/trnV*

PCR primer H1, designed by Hosaka (2002, 2003) generated three types of amplification products (a, b, c):

a. The shortest fragment of about 200 bp was detected in 36 accessions containing three accessions of the wild species *S. berthaultii*, 30 accessions of *S. tuberosum* Chilotanum group and three accessions of *S. tuberosum* Andigenum group 4x, two from Argentina (VIR 3154 from Catamarca and VIR 3231 from Jujuy Provinces) and one from Peru (VIR 8931, Junin Province) (Table 2, Supplemental file 1). This fragment corresponds to the 241 bp deletion in the *ndhC/trnV* region and to T

type of plastid DNA described by Hosaka (2003). All 36 accessions with T type plastid DNA had the same plastid SSR alleles and were characterized by haplotype #III (Table 2), and T type plastid DNA was not found outside haplotype #III. Thirty-three cultivated potato accessions with T type plastid DNA have several unique plastid SSR alleles not found in the set of the other cultivated potatoes: STCP1_117; STCP2_125, STCP3_87; and NTCP14_149) (Supplemental file 4).

b. A fragment of approximately 440 bp indicates the absence of the 241 bp deletion in the *ndhC/trnV* region (Hosaka 2003); it was detected in 356 of 392 accessions of our subset having many different plastid SSR haplotypes (Table 2).

c. The highest molecular weight fragment with approximate 500 bp size was detected in three accessions (*S. brevicaule* VIR 23061, *S. vernei* PI 230562, *S. vernei* PI 320332). All three of these accessions have plastid haplotype #XVI; all were collected in Argentina (Table 2, Supplemental file 1). The sequencing of H1 amplicons (*ndhC/trnV* region) for both *S. vernei* accessions was published by Hosaka (2003) who demonstrated the 65 bp duplication in this region. Our data (not shown) indicate identical DNA sequencing results of the *ndhC/trnV* region with the 65 bp duplication in yet another accession, *S. brevicaule* (accession VIR 23061).

Discussion

Plastid genomes in *Solanum* can only be disseminated by seeds and tubers, and in combination with nuclear markers can be very useful to reveal hybrid speciation (Wendel and Doyle 1998; Ennos et al. 1999). This is especially evident in the "bitter potatoes" *S. curtilobum*, *S. juzepczukii*, and in Chilean landraces of *S. tuberosum*.

Origin of cultivated potatoes

Origin of S. curtilobum and S. juzepczukii

Data obtained in the present study are in agreement with the earlier findings (Bukasov 1939; Hawkes 1962; Schmiediche et al. 1982) that *S. juzepczukii* (2n = 3x) originated from a cross that involved wild species *S. acaule* (2n = 4x) as the maternal parent and *S. tuberosum* Andigenum Group 2x (as *S. stenotomum*) as the paternal parent. Neighbor joining analysis revealed the close relationship between the plastid DNA type #IV of eleven accessions of *S. juzepczukii* and the related wild species *S. acaule*, *S. albicans* and *S. boliviense*. One of 12 accessions of *S. juzepczukii* atypically had haplotype #I and grouped in a Group A with accessions of *S. tuberosum* Andigenum Group, rather than grouping with the wild species putative progenitor *S. acaule* and related species of Group Ba, suggesting a possible reciprocal hybrid origin of some accessions of *S. juzepczukii* (where *S. acaule* could be pollinator). Previous plastid DNA data demonstrated common C type in *S. acaule* and in two analyzed accessions of *S. juzepczukii*, although C type was not specific for *S. acaule* and it was also detected in a members of the Andigenum group (Sukhotu et al. 2004, 2005).

Hawkes (1962) successfully resynthesized the pentaploid *S. curtilobum* by pollinating a triploid artificial *S. juzepczukii* (forming unreduced gametes) with pollen of the tetraploid cultivar 'Teton' of *S. tuberosum*. However, later Schmiediche et al. (1982) were unable to resynthesyze pentaploid hybrids using *S. juzepczukii* as maternal plants with bulked pollen from 30 tetraploid Andean landraces. Based on results of Hawkes (1962); Sukhotu et al. (2004) expected C type wild progenitor *S. acaule* be present in related cultivated species *S. curtilobum*, however both two analyzed pentaploid accessions had another S type plastid DNA that authors explained by a genetic modification during *S. curtilobum* formation. Our results shed light on the maternal lineage of *S. curtilobum*. All ten accessions of *S. curtilobum* from different geographic locations shared the same plastid SSR haplotype #I together with the majority members of the Andigenum Group, not with representatives of *S. juzepczukii* having haplotype #IV. At the same time, nuclear SSRs documented close relationship between the same *S. acaule*, *S. curtilobum*, and *S. juzepczukii* accessions (Gavrilenko et al. 2010). The results obtained here provide evidence of a maternal origin of *S. curtilobum* from members of Andigenum Group, and not from *S. juzepczukii* as was proposed before. Besides all accessions of *S. curtilobum*, haplotype #I was also detected in 82.7 % of diploid members of the Andigenum Group, in 51.9 % of triploid, and in 14.9 % of tetraploid members of the

Andigenum Group (Table 2). Accordingly we propose different scenarios for the origin of *S. curtilobum* as: Adg4x × *S. juzepczukii* (unreduced gametes); Adg2x (unreduced gametes) × *S. juzepczukii* (unreduced gametes); Adg3x (unreduced gametes) × *S. acaule*. These assumptions correlate with observations of Hawkes (1962) that mixed fields of representatives of cultivated species: *S. curtilobum*, *S. juzepczukii*, members of Andigenum group and wild species *S. acaule* growing as a weed are frequent on the altiplano.

Origin of S. ajanhuiri

Landraces of *S. ajanhuiri* are polymorphic, combining two major morphotypes 'Ajawiri' and 'Yari' (with bitter and non-bitter cultivars) that probably have different hybrid origins as proposed by Huamán et al. (1982). Earlier, Hawkes (1958) proposed an origin of *S. ajanhuiri* from crosses of *S. juzepczukii* (as maternal parent) with *S. tuberosum* Andigenum Group 2x (as *S. stenotomum*). Huamán et al. (1982) re-synthesized such crosses and obtained only a few seeds that gave aneuploid plants. Alternatively, Huamán et al. (1982) hypothesized the origin of *S. ajanhuiri* from natural crosses between *S. tuberosum* Andigenum Group 2x as the maternal parent and wild species *S. boliviense* (as *S. megistacrolobum*) as the paternal parent. He demonstrated highly viable seeds in such crosses, whereas reciprocal combinations with *S. boliviense* used as female parent gave only a few seeds with very poor germination.

Our Neighbor joining analysis placed together in group Ba: two of three accessions of *S. ajanhuiri*, 25 representatives of wild species *S. boliviense* and *S. acaule* that are closely related (Kardolus 1998; Nakagawa and Hosaka 2002), and 11 accessions of *S. juzepczukii*. The third accession of *S. ajanhuiri* has haplotype #I together with the majority of cultivated diploids, indicating possible reciprocal crosses in its origin with *S. tuberosum* Andigenum group 2x as the maternal and *S. boliviense* as the paternal parent. Nuclear SSR analysis of Spooner et al. (2007b) included the same accessions of *S. ajanhuiri* (with plastid SSR haplotype #I) into a common group with *S. acaule* and *S. juzepczukii*. Thus, our present data indicate the polymorphic nature and suggest multiple maternal origins of *S. ajanhuiri* hybrid complex. However, a much wider sampling is needed to validate our observations that are based on a three samples of *S. ajanhuiri*.

In summary, the separate grouping of most accessions of the three hybridogenic cultivated species *S. ajanhuiri*, *S. curtilobum*, *S. juzepczukii* and wild species *S. acaule* with nuclear SSR data (Spooner et al. 2007b; Gavrilenko et al. 2010) and allelic variants of these species with both the *S. tuberosum* Andigenum Group and with wild species *S. acaule* and *S. boliviense* (Rodríguez et al. 2010) support hybrid origins long proposed based on morphological, physiological and distributional data (Juzepczuk and Bukasov 1929; Juzepczuk 1937; Bukasov 1939; Hawkes 1944, 1958, 1962). Nuclear DNA analyses correlate well with our present plastid results, showing most accessions of *S. ajanhuiri* and *S. juzepczukii* in group Ba, together with their proposed wild ancestor species (Fig. 1), with the exceptions shown in our study suggesting possible multiple origins of *S. ajanhuiri* and *S. juzepczukii* from reciprocal crosses. Contradiction among nuclear SSR and plastid results in relation to *S. curtilobum* supports another maternal lineage in its origin related to Andigenum group.

S. tuberosum Chilotanum Group

We confirmed relatively low plastid diversity in Chilean landraces (members of *S. tuberosum* Chilotanum Group) previously demonstrated by Hosaka (2002) and Spooner et al. (2007b). In our study, of 34 analyzed accessions of *S. tuberosum* Chilotanum Group, two possess haplotype #V, one has a unique haplotype close to haplotype #V, one has haplotype #II, and 30 Chilean landraces share the distinct plastid SSR haplotype #III and contain the 241 bp plastid deletion corresponding to T type plastid DNA. We showed that the haplotype #III (or T type plastid DNA) is distinct not only by the presence of the 241 bp deletion, but also by many plastid SSR alleles which were not detected in the other haplotypes of cultivated species.

In addition, three accessions of *S. tuberosum* Andigenum Group 4x (two from Argentina and one from Peru) and three Argentinean accessions of wild species *S. berthaultii* (all from Province Salta, Argentina) also share haplotype #III and possess this plastid deletion. Our result supports a conclusion of Hosaka (2002, 2003, 2004) about the maternal lineage of wild species *S. berthaultii* (as *S. tarijense*) in the origin of Andean tetraploid landraces of *S. tuberosum* Andigenum

Group having the 241 bp deletion of plastid DNA and their further spread to Chile to produce *S. tuberosum* Chilotanum Group possessing T type plastid DNA. We also demonstrated that most (ten of 15 or 67 %) accessions of *S. berthaultii* without the 241 bp deletion are close to haplotype #III and occur in the same group Bb together with Chilean landraces. During spread of cultivated potatoes from the centre of diversity in the central Andes to Chile, hybridization with the other southern wild species *S. berthaultii* (with populations having haplotype #III) may have occurred. These hybrids could have been spread further south and were able to adapt to the long Chilean days. Only about 10 % of Chilean landraces possess other types of plastid DNA (different from haplotype #III or T type). Possibly, initial introductions of cultivated potatoes from southern Peru to northern Bolivia to Argentina and then to Chile (by man or seeds by birds) might have included much greater genetic diversity, but many haplotypes were eliminated due to their inability to adapt to the new environments, or they were lost because of their low frequencies in the primary introductions.

Spooner et al. (2012) showed that the wild Chilean species *S. maglia* grouped with Chilean landraces with nuclear SSRs, but not with the 241 bp plastid deletion, providing conflicting and ambiguous evidence on the involvement of *S. maglia* in the origin of *S. tuberosum* Chilotanum Group. In our study only one of 34 accessions of *S. tuberosum* Chilotanum Group had the same haplotype #II together with all three accessions of *S. maglia*. Based on results of Spooner et al. (2012) and present data we can suggest participation of *S. maglia* in the formation Chilean tetraploid landraces, possibly mainly as paternal contributors.

Origin of S. tuberosum Andigenum group

Morphological (Van den Berg et al. 1998) and AFLP (Spooner et al. 2005) data are concordant in demonstrating *S. candolleanum* (or a closely related or extinct species) in domestication of members of *S. tuberosum* Andigenum group. At least for the diploid members of the Andigenum Group, this matches hypotheses with plastid DNA data of Sukhotu et al. (2004, 2006), and Sukhotu and Hosaka (2006) who grouped several accessions of *S. candolleanum* (as *S. bukasovii*, *S. candolleanum*, *S. multidissectum*), all possessing the same plastid type S with members of the *S. tuberosum* Andigenum group.

In our study almost all members of *S. tuberosum* Andigenum group having predominant haplotypes #I and #II were in group A, together with most (63.2 %) representatives of *S. candolleanum* and few (5.4 %) representatives of *S. brevicaule* (former northern and southern members of *S. brevicaule* complex). The presence of haplotype #I in the majority of diploid members of *S. tuberosum* Andigenum group and in two wild species accessions of *S. candolleanum*, *S. brevicaule* could be interpreted to represent a common origin and ancestral haplotype for cultivated species. However haplotype #I also includes one accession of *S. tarijense* (PI 275154). Finding the one accession of Argentinean species *S. tarijense* in the group of haplotype #I is unexpected, but this is not a chance. As we mentioned above, haplotype #I is easily identified using the NTCP6 primer pairs that distinguish the allele NTCP6 127 bp corresponding to a 48 bp deletion in the amplified *rps16/trnQ* region (Hosaka 2003). Hosaka (2003) found two of 62 *S. berthaultii* accessions with the same NTCP6_127 bp fragment which is typical for S type plastid DNA in his study. Later, we screened 83 accessions of *S. berthaultii* for the presence of allele NTCP6_127 bp and found one additional sample (PI 473240) with this allele (NTCP6_127), typical for haplotype #I in the present study. This finding could indicate gene flow between wild species *S. berthaultii* (as pollen donor) and populations of *S. brevicaule* (southern members of *S. brevicaule* complex) having haplotype #I.

Genetic diversity in cultivated potatoes

We found very restricted plastid SSR diversity in cultivated species in comparison with closely related wild species. Fully 206 of 237 (or 88 %) of the cultivated species accessions shared four most frequent haplotypes: #I, #II, #III and #IV. There was much greater diversity in the 155 wild species accessions, which consisted of 110 haplotypes, dramatically highlighting decreasing levels of genetic diversity through domestication. Our data suggest that cultivated species originated from a limited populations of wild species. For example, the pentaploid cultivated species *S. curtilobum* is monomorphic in its plastid haplotype #I. Similarly, *S. juzepczukii* has a greatly reduced plastid diversity, in contrast to its wild

species introgressant *S. acaule* (Table 2). Similarly, we documented low plastid diversity in *S. tuberosum* Chilotanum Group ($h = 0.223$), whereas its possible plastid donor wild species *S. berthaultii* is highly heterogeneous. Fully 100 of 104 (96 %) of *S. tuberosum* Andigenum Group 2x shared two main plastid haplotypes #I and #II, that explained their low heterogeneity ($h = 0.296$). In contrast, plastid diversity of the 93 accessions of their wild species progenitors or close relatives (*S. brevicaule*, *S. candolleanum*) exhibited much higher diversity, divided into 69 haplotypes (Table 2). Sukhotu and Hosaka (2006) and Sukhotu et al. (2006) demonstrated limited genetic diversity in diploids of *S. tuberosum* Andigenum group having mainly two major plastid DNA types (S and A1 in their study), and suggesting that *S. tuberosum* Andigenum group 2x either (a) had dual origins by successive domestication from two different wild species or (b) its diversity resulted from introgression after it first arose.

Our data are in general agreement with studies of Sukhotu et al. (2005) and Sukhotu and Hosaka (2006) who showed limited diversity of *S. tuberosum* Andigenum group 2x in comparison with *S. tuberosum* Andigenum group 4x, and support earlier hypotheses that tetraploids of *S. tuberosum* Andigenum group arose from diploid Andean landraces by sexual polyploidization and introgression from related wild species. In our study *S. tuberosum* Andigenum group 4x is the most heterogeneous in plastid DNA ($h = 0.789$; 13 haplotypes distributed in all groups, Fig. 1) among all cultivated species. This might be explained by their hybridization (as paternal parents) with representatives of different wild species (as maternal parents) in different locations. Some of the tetraploid Andean landraces from the northern Andes have haplotypes #VI and #VIII which were not found in the southern Andes and contrastingly, some of *S. tuberosum* Andigenum group 4x accessions with haplotypes #V and #III were detected mainly in the southern Andes and were absent farther north.

We demonstrated relatively higher plastid diversity ($h = 0.562$) in *S. tuberosum* Andigenum group 3x (as *S. chaucha*) in comparison with *S. tuberosum* Andigenum group 2x that could be explained by a different origin of these triploid landraces, both through meiotic polyploidization of cultivated diploids and through crosses of diverse *S. tuberosum* Andigenum group 4x as maternal parent with less polymorphic *S. tuberosum* Andigenum group 2x. These triploid Andean landraces having haplotypes #I and #II are divided into two morphological groups (Supplemental file 9), possibly because of sterility that prevent gene flow in contrast to diploid landraces with free reciprocal crossability. The existence of distinct morphotypes in *S. tuberosum* Andigenum group 3x in their tuber protein profiles was described earlier by Jackson et al. (1977).

We conclude that data obtained with different germplasm collections in the present study and in previous plastid DNA studies of Hosaka with collaborators are, in general, in agreement with each other, although they used very restricted subset of *S. tuberosum* Andigenum group 3x, *S. curtilobum*, and *S. juzepczukii*, and another plastid DNA marker set (Hosaka 2002, 2003; Sukhotu et al. 2004, 2005, 2006; Sukhotu and Hosaka 2006).

Geographical distribution of plastid DNA haplotypes of cultivated species

It generally has been accepted that the center of domestication and genetic diversity of cultivated potatoes lies in the central Andes where all ploidy levels of cultivated species have a high varietal diversity. Sukhotu et al. (2005) using another set of plastid SSR markers found greatest diversity of tetraploid Andean landraces in the central Andes, particularly in Bolivia. We documented high plastid DNA diversity in the complex of all cultivated potatoes in the central Andes, lower in the northern Andes (Ecuador, Columbia), and lowest in Chile.

Our present results can generally discriminate three major gene pools within plastid DNA in cultivated potato species and underline three different maternal lineages in their origin.

1. Haplotypes genetically clustered into the Group A, which are common for the diploid, triploid and tetraploid members of Andigenum group and most frequent in the central Andes at latitude −10 – −15. This gene pool (combining two main plastid haplotypes #I and #II, small haplotypes #VI and #VIII, and six unique haplotypes) represents probably the earliest potato domestication events. The plastid markers could not separate members of *S. tuberosum* Andigenum group 2x, 3x, 4x as their accessions share the same two predominant haplotypes (#I and #II) with a

few exceptions. Based on our present data and prior molecular results we conclude that *S. tuberosum* Andigenum group 2x, 3x, and 4x cytotypes cannot reliably be divided into distinct species but are better treated as a highly polymorphic Andigenum Group as outlined in Spooner et al. (2007b) and Ovchinnikova et al. (2011).

2. Haplotypes clustered in group Ba are common for hybridogenic highland cultivated species *S. juzepczukii* (and probably for *S. ajanhuiri*) that had a maternal lineage related to the wild species progenitors present in Group Ba.
3. Haplotypes clustered in group Bb, which combine Chilean and Argentinean representatives having haplotype #III or T type plastid DNA which found mainly in the *S. tuberosum* Chilotanum group and genetically related haplotypes (without 241 bp deletion) which all are clustered in group Bb. These haplotypes derived from the maternal lineage related with Argentinean wild species *S. berthaultii* or their ancestor species.

Potato domestication involved several hybridization events in different periods. It is unclear if the plastid haplotypes in the *S. tuberosum* Andigenum group represent original lineages or later introgression events subsequent to a single origin of this group in the Andes, as supported by AFLP data (Spooner et al. 2005). If the plastid DNA track primary evolutionary events, they support an original domestication from limited populations of Andean diploid wild species *S. candolleanum* or *S. brevicaule* having haplotypes #I, which led to formation of diploid landraces of *S. tuberosum* Andigenum group in southern Peru—northern Bolivia. This supports the restricted origin hypothesis. However, after potatoes were already domesticated in the central Andes, interspecific hybridization among cultivated species and genotypes of different wild gene pools has played an important role in the broadening diversity of primary domesticated forms. From the original domestication event(s) which led to formation of the *S. tuberosum* Andigenum group, secondary sites of domestication occurred that allowed cultivated species to be spread in the widest ecological and geographical conditions in the highland Andes (due to likely introgression from wild species *S. acaule* and *S. boliviense*) and in lowland southern coast Chile (as result of possible introgression from maternal lineage of *S. berthaultii*, and of possible introgression from paternal lineage of *S. maglia*).

Acknowledgments This study was supported by International Science and Technology Center (ISTC), grant no. 3329, by the ARS/USDA (http://www.vir.nw.ru/biot), by RFBR grant no. 11-04-01002-a and by NSF DEB 0316614 entitled PBI *Solanum*: A worldwide treatment http://www.nhm.ac.uk/research-curation/research/projects/solanaceaesource/). We also acknowledge Dr. S. Kiru for providing germplasm from VIR potato collection.

References

Brown CR (1999) A native American technology transfer: the diffusion of potato. HortScience 34:817–821

Bryan GJ, McNicoll J, Ramsay G, Meyer RC, DeJong WS (1999) Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. Theor Appl Genet 99:859–867

Bukasov SM (1933) The potatoes of South America and their breeding possibilities. Trudy po Prikladnoj Botanike Genetike i Selekcii, Supplement 58:1–192

Bukasov SM (1939) The origin of potato species. Physis (Buenos Aires) 18:41–46

Bukasov SM (1966) Die Kulturarten der Kartoffel und ihre wildwachsenden Vorfahren. Zeitschr für Pflanzenzüchtung Bd 55. Hft 2:139–164

Bukasov SM (1978) Systematics of the potato. Trudy Prikl Bot Gen Selekcii 62:3–35

Chung HJ, Jung JD, Park HW, Kim JH, Cha HW, Min SR, Jeong WJ, Liu J (2006) The complete chloroplast genome sequences of *Solanum tuberosum* and comparative analysis with Solanaceae species identified the presence of a 241-bp deletion in cultivated potato chloroplast DNA sequence. Plant Cell Rep 25:1369–1379

Darwin C (1845) A naturalist's voyage. Journal of research into the natural history and geology of the countries visited during the voyage of H.M.C. Beagle round the world, John Murray, London

Ennos RA, Sinclair WT, Hu XS, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: Hollingsworth PM, Bateman RM, Gornall RJ (eds) Molecular systematics and plant evolution. Taylor & Francis, London, pp 1–19

Gavrilenko T, Antonova O, Ovchinnikova A, Novikova L, Krilova E, Mironenko N, Pendinen G, Islamshina A, Shvachko N, Kiru S, Kostina L, Afanasenko O, Spooner DM (2010) A microsatellite and morphological assessment of the Russian National cultivated potato collection. Genet Resour Crop Evol 57:1151–1164

Grun P (1973) Cytoplasmic sterilities that separate the Group *Tuberosum* common potato from its putative tetraploid ancestor. Evolution 27:633–643

Grun P (1990) The evolution of cultivated potatoes. In: Bretting PK (ed) New perspectives on the origin and evolution of New World domesticated plants. Econ Bot 44(3 Suppl): 39–55

Grun P, Ochoa C, Capage D (1977) Evolution of cytoplasmic factors in tetraploid common potatoes. Amer J Bot 64:412–420

Hawkes JG (1944) Potato collecting expeditions in Mexico and South America. II. Systematic classification of the collections. Bull Imp Bur Plant Breed and Genet, Cambridge

Hawkes JG (1956) Taxonomic studies on the tuber-bearing *Solanums*. 1. *Solanum tuberosum* and the tetraploid complex. Proc Linn Soc Lond 166:97–144

Hawkes JG (1958) Kartoffel. 1. Taxonomy, cytology and crossability. In: Kappert H, Rudorf W (ed) Handbuch der Pflanzenzüchtung. Paul Parey, Berlin 3:1–43

Hawkes JG (1962) The origin of *Solanum juzepczukii* Buk. and *S. curtilobum* Juz. et Buk. Z Pflanzenz 47:1–14

Hawkes JG (1990) The potato: evolution, biodiversity, and genetic resources. Smithsonian Institution Press, Washington

Hijmans RJ, Spooner DM, Salas A, Guarino L, De La Cruz J (2002) Atlas of wild potatoes. Systematic and Ecogeographic Studies of Crop Genepools 10:1–129

Hosaka K (1986) Who is the mother of the potato?—Restriction endonuclease analysis of chloroplast DNA of cultivated potatoes. Theor Appl Genet 72:606–618

Hosaka K (1995) Successive domestication and evolution of the Andean potatoes as revealed by chloroplast DNA restriction endonuclease analysis. Theor Appl Genet 90:356–363

Hosaka K (2002) Distribution of the 241 bp deletion of chloroplast DNA in wild potato species. Amer J Potato Res 79:119–123

Hosaka K (2003) T-type chloroplast DNA in *Solanum tuberosum* L. ssp. *tuberosum* was conferred from some populations of *S. tarijense* Hawkes. Amer J Potato Res 80:21–32

Hosaka K (2004) Evolutionary pathway of T-type chloroplast DNA in potato. Amer J Potato Res 81:153–158

Hosaka K, Hanneman RE Jr (1988a) The origin of the cultivated tetraploid potato based on chloroplast DNA. Theor Appl Genet 76:172–176

Hosaka K, Hanneman RE Jr (1988b) Origin of chloroplast DNA diversity in the Andean potatoes. Theor Appl Genet 76:333–340

Hosaka K, Ogihara Y, Matsubayashi M, Tsunewaki K (1984) Phylogenetic relationship between the tuberous *Solanum* species as revealed by restriction endonuclease analysis of chloroplast DNA. Jpn J Genet 59:349–369

Huamán Z, Spooner DM (2002) Reclassification of landrace populations of cultivated potatoes (*Solanum* sect. *Petota*). Amer J Bot 89:947–965

Huamán Z, Hawkes JG, Rowe PR (1982) A biosystematic study of the origin of the diploid potato, *Solanum ajanhuiri*. Euphytica 31:665–675

Huanco V (1991) Potencial de las papas amargas en el altiplano de Puno. Perú. In: Rea J, Vacher JJ (eds) La papa amarga. I Mesa Redonda, Peru, pp 25–26

Jackson MT, Hawkes JG, Rowe PR (1977) The nature of *Solanum* × *chaucha* Juz. et Buk., a triploid cultivated potato of the South American Andes. Euphytica 26:775–783

Juzepczuk SW (1937) New species of the genus *Solanum* L. in the group *Tuberarium* Dun. Bull Acad Sci USSR 2:295–331

Juzepczuk SW, Bukasov SM (1929) A contribution to the question of the origin of the potato. Trudy Vsesoyuzn Szeda Genet Selek 3:593–611

Kardolus JP (1998) A biosystematic analysis of *Solanum acaule*. Dissertation, Wageningen University and Research Center Publications

Kawagoe Y, Kikuta Y (1991) Chloroplast DNA evolution in potato (*Solanum tuberosum* L). Theor Appl Genet 81:13–20

Miller JT, Spooner DM (1999) Collapse of species boundaries in the wild potato *Solanum brevicaule* complex (Solanaceae, *S.* sect. *Petota*): molecular data. Plant Syst Evol 214:103–130

Nakagawa K, Hosaka K (2002) Species relationships between a wild tetraploid potato species, *Solanum acaule* Bitter, and its related species as revealed by RFLPs of chloroplast and nuclear DNA. Am J Potato Res 79:85–98

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Ovchinnikova A, Krylova E, Gavrilenko T, Smekalova T, Zhuk M, Knapp S, Spooner DM (2011) Taxonomy of cultivated potatoes (*Solanum* section *Petota*: Solanaceae). Bot J Linn Soc 165:107–155

Provan J, Powell W, Dewar H, Bryan G, Machray GC, Waugh R (1999) An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity. Proc R Soc Lond B Biol Sci 266:633–639

Raker C, Spooner DM (2002) The Chilean tetraploid cultivated potato, *Solanum tuberosum*, is distinct from the Andean populations; microsatellite data. Crop Sci 42:1451–1458

Rodríguez F, Ghislain M, Clausen AM, Jansky SH, Spooner DM (2010) Hybrid origins of cultivated potatoes. Theor Appl Genet 121:1187–1198

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Salaman RN (1946) The early European potato: its character and place of origin. J Linn Soc Lond Bot 53:1–27

Schmiediche PE, Hawkes JG, Ochoa CM (1982) Breeding of the cultivated potato species *Solanum juzepczukii* Buk. and *S. curtilobum* Juz. et Buk. II. Euphytica 31:695–707

Simmonds NW (1964) Studies of the tetraploid potatoes. 1. The variability of the Andigena Group. J Linn Soc Lond Bot 58:461–474

Spooner DM (2009) DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. Amer J Bot 96:1177–1189

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus AFLP genotyping. Proc Natl Acad Sci USA 102:14694–14699

Spooner DM, Fajardo D, Bryan GJ (2007a) Species limits of *Solanum berthaultii* Hawkes and *S. tarijense* Hawkes and the implications for species boundaries in *Solanum* sect. Petota Taxon 56:987–999

Spooner DM, Núñez J, Trujillo G, Del Rosario Herrera M, Guzmán F, Ghislain M (2007b) Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. Proc Natl Acad Sci USA 104:19398–19403

Spooner DM, Gavrilenko T, Jansky SH, Ovchinnikova A, Krylova E, Knapp S, Simon R (2010) Ecogeography of ploidy variation in cultivated potato (*Solanum* sect. *Petota*). Amer J Bot 97:2049–2060

Spooner DM, Jansky S, Clausen A, Del Rosario Herrera M, Ghislain M (2012) The enigma of *Solanum maglia* in the origin of the Chilean cultivated potato, *Solanum tuberosum* Chilotanum Group. Econ Bot 66:12–21

Sukhotu T, Hosaka K (2006) Origin and evolution of Andigena potatoes revealed by chloroplast and nuclear DNA markers. Genome 49:636–647

Sukhotu T, Kamijima O, Hosaka K (2004) Nuclear and chloroplast DNA differentiation in Andean potatoes. Genome 47:46–56

Sukhotu T, Kamijima O, Hosaka K (2005) Genetic diversity of the Andean tetraploid cultivated potato (*Solanum tuberosum* L. subsp. *andigena* Hawkes) evaluated by chloroplast and nuclear DNA markers. Genome 48:55–64

Sukhotu T, Kamijima O, Hosaka K (2006) Chloroplast DNA variation in the most primitive cultivated diploid potato species *Solanum stenotomum* Juz. et Buk. and its putative wild ancestral species using high-resolution markers. Genet Res Crop Evol 53:53–63

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molec Biol Evol 24:1596–1599

Ugent D (1970) The potato: what is the origin of this important crop plant, and how did it first become domesticated? Science 170:1161–1166

Ugent D, Dillehay T, Ramirez C (1987) Potato remains from a late Pleistocene settlement in South central Chile. Econ Bot 41:17–27

Van den Berg RG, Miller JT, Ugarte ML, Kardolus JP, Villand J, Nienhuis J, Spooner DM (1998) Collapse of morphological species in the wild potato *Solanum brevicaule* complex (sect. *Petota*). Amer J Bot 85:92–109

Wendel JE, Doyle JJ (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis PS, Doyle JJ (eds) Plant molecular systematics II. Kluwer Academic Publishers, Dordrecht, pp 265–296